Information Sciences 290 (2015) 7–21

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Predicting electrical evoked potential in optic nerve visual prostheses by using support vector regression and case-based prediction

Jie Hu [a,*], Jin Qi [a], Yinghong Peng [a], Qiushi Ren [b]

[a] Institute of Knowledge Based Engineering, School of Mechanical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China
[b] Department of Biomedical Engineering, Peking University, 60 Yannan Yuan, Beijing 100871, China

ARTICLE INFO

Article history:
Received 24 January 2013
Received in revised form 31 July 2014
Accepted 13 August 2014
Available online 23 August 2014

Keywords:
Electrical evoked potential
Time series prediction
Feature integration
Case based prediction
Support vector regression

ABSTRACT

Electrical evoked potential (EEP) forecasting is an intelligent time series prediction (TSP) activity to explore the temporal properties of electrically elicited responses of the visual cortex triggered by various electrical stimulations. Our previous studies used support vector regression (SVR) as a TSP predictor to forecast temporal EEP values. SVR shows high prediction performance but with high computation time for multivariable stimulation inputs in EEP prediction. To reduce the computational burden of SVR and further improve the performance, this paper utilizes technique of case-based prediction (CBP) to integrate the initial stimulation variables into an integrated stimulation value (ISV), and total four independent CBPs are used to achieve the stimulation feature integration. Then the temporal samples are extracted from transformed data to construct a new SVR regression model to perform the prediction activity. The new hybridizing system is named as CBSVR, which was also empirically tested with data collected from actual EEP electrophysiological experiments. Both 30-fold cross-validation method and adapted point predictive accuracy (PPA) index were used to compare the predictive performances between CBSVR, classical CBP approaches, single SVR model and other common TSP methods. Empirical comparison results show that CBSVR is feasible and validated for EEP prediction in visual prostheses research.

1. Introduction

The electrical evoked potential (EEP) forecasting problem arises out of optic nerve visual prostheses research whose purpose is to deliver the electrical impulses to optic nerve with the penetrating electrode, and the visual perception is then emerged from the visual cortex [39]. Therefore, it is necessary to investigate the temporal properties of electrically elicited responses of the visual cortex triggered by optic nerve stimulation, and EEP data should be collected to indicate the effects of different stimulation parameters [31,38]. However, subjected to experiment cost and material restriction, the collected EEP data are too sparse and inadequate to be analyzed in optic nerve visual prosthesis research [32]. The motivation of EEP prediction is to explore another way to obtain the new EEP behaviors instead of performing the more expensive electrophysiological experiments, that is, employ an intelligent approach to predict the new EEP movements for new electrical

* Corresponding author. Tel.: +86 21 34206552; fax: +86 21 34206313.
E-mail address: hujie@sjtu.edu.cn (J. Hu).

http://dx.doi.org/10.1016/j.ins.2014.08.036

stimulations, according to the existing EEP elicited data. As the characteristics of EEP elicited data are time series, inherently noisy, non-stationary, and deterministically chaotic, this means that not only is a single data series non-stationary in the sense of the mean and variance of the series, but the relationship of the data series to other related data series may also be changing [24].

Generally, time series can be modeled by linear and non-linear models. The linear models such as autoregressive integrated moving average (ARIMA) [11], multiple discriminant analysis (MDA) [27] and exponential smoothing [28] have been extensively investigated, but they only work well when the data dependency is linear. The non-linear models used for time series prediction (TSP) were chiefly the neural networks (NNs), e.g., back-propagation neural network [3,13], multi-layer feed-forward neural network [15], sparsely connected neural networks [4,16], Kalman filter-based Elman network [43], and hybrid neural networks [1,28,41]. Although these networks work well, they have some inherent drawbacks, such as the problem of multiple local minima, the choice of number of hidden units and the danger of overfitting [31]. Meanwhile, case based prediction (CBP) originated from case based reasoning methodology began to be employed in this area as well [19,20,31], which allows experts to make their decision and predict the future behavior of the current time series by referring similar old case from past [31]. So it dose not require any training data distribution assumption for input cases (linear or non-linear) and easily being understood. But CBP approach also often neglects or oversimplifies the time-varying information hidden in temporal data when it is used to predict temporal behavior. Recently, support vector regression (SVR), based on the unique theory of the structural risk minimization principle of support vector machine (SVM), has been provided as a novel approach to perform TSP [5,22] and abundant researches have proved SVR can achieve better generalization ability than traditional NNs [2,21,29,30,36]. Another advantage of the SVR approach is the applicability of the kernel-trick, which makes it more practicable [42].

Accordingly, we have utilized SVR to predict EEP values in our previous studies [29,30] where stimulation variables and corresponding temporal EEP values were considered as inputs and outputs of SVR model directly. However, the potential weakness of SVR in EEP prediction is its relative high computation time, since training an SVR model is equivalent to solving a linearly constrained quadratic programming problem. Suppose that EEP time series data derive from m EEP electrophysiological experiments, and each experiment contains n time points and q stimulation variables, then the size of EEP prediction problem is $O(n \times m \times q)$. Traditionally, feature reduction (FR) is a chief way to reduce the computational burden of algorithm. However subjected to experiment cost and material restriction, the original EEP data exist in very high dimension space, which makes the utilization of classical FR methods may confront small sample size (SSS) problem. So this paper attempts to propose another FR way to reduce the computational complexity of SVR and further improve the predictive performance. The contribution of this paper is twofold.

On the one hand, we try to take the advantage of CBP's relatively simplicity for feature-oriented computation and SVR's high generalization for TSP to predict EEP values. This is done by implementing the technology of CBP to integrate q stimulation variables into an integrated stimulation value (ISV) as the univariate input of SVR, and constructing a new SVR-based prediction model with these transformed data. So this new EEP predictor is named as CBSVR, whose size decreases to $O(n \times m)$. On the other hand, we also try to investigate the superiority of the new predictor in actual EEP prediction. The predictive performances of hybridized CBSVR model and other classical TSP methods such as CBP, single SVR and NN are planned to be compared. Empirical EEP data are collected from the experiments of rabbit's visual cortex responses for optic nerve stimulation carried out inside the pia mater, on the pia mater and on the dura mater. In the assessment of predictive performances, we employ 30-time cross-validation strategy by combining hold-out method and leave-one-out cross-validation, as Li and Sun [20] did, then statistical analysis is employed to find whether or not there are significant differences among comparative predictors in the light of the predictive results.

This paper is organized as follows. Section 2 makes a brief review on related methods. Section 3 addresses how to build the construction of CBSVR model for EEP prediction. Section 4 designs the predictive experiment, while the empirical results and analysis are presented in Section 5. Section 6 draws some conclusions.

2. Research background

2.1. Support vector regression for time series prediction

The basic idea of SVR for TSP is to map the temporal data into a higher dimension space via a nonlinear mapping and then to do linear regression in this space. Therefore, the regression approximation addresses the problem of estimating a function based on a given time-stamp data set. Given the set of raw training data $\{\langle x_i, y_i\rangle\}(i = 1, 2, \ldots, p)$, in which i represents the data sequence, i.e., the ith time node, $x \in R^q$ is a q-dimensional input vector of SVR, $y \in R$ is the corresponding value, and p is the total number of time nodes. Then, SVR model approximates the unknown function with form $f(x) = \omega^T\varphi(x) + \beta$, where $\varphi(x)$ is the high dimensional feature space, nonlinearly mapped from the input space, ω represents the normal vector and β is the bias. By introducing the ε-insensitive loss function, SVR approximates the linear $f(x)$ satisfies the following conditions:

$$\begin{aligned} &|y_i - f(x_i)| = |y_i - \omega^T\varphi(x_i) - \beta| \leqslant \varepsilon \\ &x_i,\ \omega \in R^q \quad y_i,\ \beta \in R \end{aligned} \tag{1}$$

To guarantee the satisfaction of the aforesaid conditions, coefficients ω and β are optimized by solving: